END
DATE
FILMED
4—80
DDC

MICROCOPY RESOLUTION TEST CHART
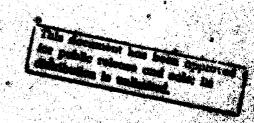NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL

ANY ITERATION FOR POLYNOMIAL EQUATIONS
USING LINEAR INFORMATION HAS INFINITE COMPLEXITY

G. W. Wasilkowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(On leave from the University of Warsaw)

September 1979

DTIC
SELECTE
MAR 6 1980

C

DEPARTMENT
of
COMPUTER SCIENCE

Carnegie-Mellon University

80 3 6 005

ANY ITERATION FOR POLYNOMIAL EQUATIONS
USING LINEAR INFORMATION HAS INFINITE COMPLEXITY .

G. W. Wasilkowski

Department of Computer Science
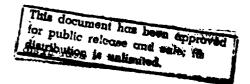Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(On leave from the University of Warsaw)

September 1979

ABSTRACT

This is the third paper in which we study iterations using linear information for the solution of nonlinear equations. ~~In Wasilkowski [78]~~ ~~and [79]~~ we have considered the existence of globally convergent iterations for the class of analytic functions. Here we study the complexity of such iterations. We prove that even for the class of scalar complex polynomials with simple zeros, any iteration using arbitrary linear information has infinite complexity. More precisely, we show that for any iteration ϕ and any integer k, there exists a complex polynomial f with all simple zeros such that the first k approximations produced by ϕ do not approximate any solution of f = 0 better than a starting approximation x_0. This holds even if the distance between x_0 and the nearest solution of f = 0 is arbitrarily small.

1. INTRODUCTION

In this paper we continue the study of iterations using linear information for the solution of nonlinear equations $f = 0$. In Wasilkowski [78] we have proven that no _stationary_ iteration using linear information can be _globally_ convergent for the class of scalar analytic functions with simple zeros. In Wasilkowski [79] we have exhibited _nonstationary_ iterations which are _globally_ convergent for the class of analytic functions with simple zeros even for the abstract case.

In this paper we deal with the complexity of iterations using linear information. We prove the surprising result that any such iteration has _infinite_ complexity even for the class $\mathfrak{J}$ of scalar complex polynomials with simple zeros. To make this negative result as strong as possible we have chosen a relatively simple class $\mathfrak{J}$. Furthermore we deal with a very general definition of information and iteration. Namely, any sequence of linear finite dimensional operators is considered as possible information, and any sequence of functionals as an iteration. We also do not specify which zero of f is approximated, and the assumptions concerning the starting points are very weak. Under these assumptions we prove that for any positive L, any integer k, and any iteration $\overline{\varphi}$ using linear information, there exists a complex polynomial f having only simple zeros such that the distance between a starting approximation x_0 and a nearest zero α of f is no larger than L and the first k approximation produced by $\overline{\varphi}$ do not approximate any zero of f better than x_0. Note that L can be arbitrarily small which means that x_0 can be arbitrarily close to α.

Let $S(f)$ denote the set of all zeros of f. By the complexity of an iteration we mean the total cost of producing an approximation x_N where N is the minimal index such that $dist(x_N, S(f)) \leq \epsilon\ dist(x_0, S(f))$ for a given number ϵ, $\epsilon < 1$. We do not specify exactly what we mean by the "cost". We merely assume that the cost of the assignment operation is not zero. Thus, the complexity of an iteration is at least proportional to N. Since we shall show that N can be arbitrarily large for some polynomials, this proves that for every ϵ the complexity is infinite in the class $\mathfrak{I}$. This is a very strong result since even assuming (theoretically) that all operations except assignment are free, the complexity is still infinite.

This paper also illustrates the important difference between the concepts of global convergence and complexity. The class of all linear information supplies enough knowledge about f to guarantee the existence of globally convergent iterations but the complexity of any such iteration is infinite.

We summarize the contents of the paper. For the reader's convenience, in Sections 2 through 5 we deal only with iterations without memory. In Sections 2 and 3 we define <u>information</u>, <u>iteration without memory</u>, <u>globally convergent</u> iteration, and <u>complexity</u> of an iteration. In Section 4 we prove two theorems which play an essential role in the proof of the main result which is established in Section 5. In Section 6 we extend all results to iterations with memory. In Section 7 we pose some open problems.

2. INFORMATION AND ITERATIONS WITHOUT MEMORY

For the reader's convenience we repeat the very general definition of information and iteration without memory introduced in Wasilkowski [79]. For simplicity, in Sections 2 through 5 we deal only with iterations without memory. The extension to the general case is given in Section 6.

Let H be the class of all complex polynomials and $\mathfrak{J}$ be the subset of H which consists of all polynomials having only simple zeros. Let $S(f)$ denote the set of all zeros of f, $f \in H$. Consider the solution of a nonlinear equation

$$(2.1) \quad f(x) = 0, \ f \in \mathfrak{J}.$$

To solve (2.1) iteratively we must know something about f. Let $L_i : H \times \mathbb{C} \to \mathbb{C}$ be a functional which is linear with respect to the first argument, i.e., $L_i(c_1 f_1 + c_2 f_2, x) = c_1 L_i(f_1, x) + c_2 L_i(f_2, x)$, $i = 1, 2, \ldots, n$. Then the <u>linear information operator</u> $\mathfrak{N}$, $\mathfrak{N} = [L_1, L_2, \ldots, L_n] : H \times \mathbb{C} \to \mathbb{C}^n$, is defined as

$$(2.2) \quad \mathfrak{N}(f,x) = [L_1(f,z_1), L_2(f,z_2), \ldots, L_n(f,z_n)], \ \forall f \in H, \ \forall x \in \mathbb{C}$$

where $z_1 = x$ and

$$z_j = \xi_j(z_1; \ L_1(f,z_1), L_2(f,z_2), \ldots, L_{j-1}(f,z_{j-1}))$$

for some functions ξ_j, $j = 2, 3, \ldots, n$. Thus any z_j depends on the previously computed information. For brevity we shall sometimes write $z_j = z_j(f)$. Let Ψ_n be the class of all such information operators.

Consider a sequence of linear information operators $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$, $\mathfrak{N}_i \in \psi_{n_i}$. Let x_0 be an approximation of a solution of (2.1). Suppose we construct a sequence of approximations $\{x_i\}$ by the formula

$$(2.3) \quad x_i = \varphi_i(x_0; \mathfrak{N}_i(f, x_0))$$

where $\varphi_i : D_{\varphi_i} \subset \mathbb{C}^{1+n_i} \to \mathbb{C}$ are functionals, $\varphi_i \in \mathfrak{I}(\mathfrak{N}_i)$. Then the sequence $\overline{\varphi} = \{\varphi_i\}$ is called an <u>iteration without memory</u> using $\overline{\mathfrak{N}}$, $\overline{\varphi} \in \overline{\mathfrak{I}}(\overline{\mathfrak{N}})$.

3. COMPLEXITY OF ITERATIONS

In this section we define the complexity of an iteration. Let

$$\text{dist}(x, S(f)) = \inf_{\alpha \in S(f)} |x - \alpha|$$

denote the distance between the point x and the set $S(f)$. Let L be a positive number and let $\overline{\varphi}$ be an iteration without memory. For any $f \in \mathfrak{F}$ and x_0 such that

(3.1) $\text{dist}(x_0, S(f)) < L,$

consider the sequence $\{x_i\}$ generated by $\overline{\varphi}$. For any ϵ, $\epsilon < 1$, define $N = N(\overline{\varphi}, \epsilon, x_0, f)$ as the minimal integer, if it exists, such that

(3.2) $\text{dist}(x_N, S(f)) \leq \epsilon \, \text{dist}(x_0, S(f)),$

and $N = +\infty$ otherwise. The number N is determined by how many iterative steps are necessary to reduce the starting error by ϵ.

Let $\text{comp}(\overline{\varphi}, \epsilon, x_0, f)$ be the total cost of computing x_N satisfying (3.2). We do not specify exactly what we mean by the "cost". We merely assume that the cost of the assignment operation is not zero. Since any iterative step performs at least one assignment operation, there exists a positive number c such that

(3.3) $\text{comp}(\overline{\varphi}, \epsilon, x_0, f) \geq cN(\overline{\varphi}, \epsilon, x_0, f), \ \forall \overline{\varphi}, \epsilon, x_0, f.$

In Wasilkowski [79] we showed there exist <u>globally convergent iterations</u>, ie., iterations which for any x_0 and f satisfying (3.1) construct a sequence $\{x_i\}$ such that

(3.4) $\lim_{i \to \infty} x_i \in S(f)$.

(This also holds for $L = +\infty$.)   Note that for any globally convergent iteration $\bar{\varphi}$, the number $N(\bar{\varphi}, \varepsilon, x_0, f)$ is finite for any positive ε, any x_0, and any fixed f from $\mathfrak{J}$.

We shall show that $N(\bar{\varphi}, \varepsilon, x_0, f)$ is unbounded for a subset of $\mathfrak{J}$.   Let

(3.5) $\mathfrak{J}(x_0) = \{f \in \mathfrak{J} : dist(x_0, S(f)) < L\}$.

Thus, $\mathfrak{J}(x_0)$ is the set of all polynomials f from $\mathfrak{J}$ for which the distance between the initial approximation x_0 and the nearest zero of f is less than L. Let

(3.6) $N(\bar{\varphi}, \varepsilon, x_0) = \sup_{f \in \mathfrak{J}(x_0)} N(\bar{\varphi}, \varepsilon, x_0, f)$

be the minimal number of iterative steps which are necessary to reduce the starting error by ε for all f from $\mathfrak{J}(x_0)$.   Similarly, let

(3.7) $comp(\bar{\varphi}, \varepsilon, x_0) = \sup_{f \in \mathfrak{J}(x_0)} comp(\bar{\varphi}, \varepsilon, x_0, f)$.

Due to (3.3),

(3.8) $comp(\bar{\varphi}, \varepsilon, x_0) \geq c N(\bar{\varphi}, \varepsilon, x_0)$, $\forall \bar{\varphi}, \varepsilon, x_0$.

It is intuitively obvious that for $\varepsilon = 0$, $N(\bar{\varphi}, 0, x_0) = +\infty$.   In Section 5 we prove that for any $L \in (0, +\infty]$, any iteration $\bar{\varphi}$ using linear information and any $x_0 \in \mathbb{C}$,

$$N(\bar{\varphi}, \varepsilon, x_0) = +\infty, \ \forall \varepsilon \in [0, 1)$$

which, due to (3.8), implies that

$$\text{comp}(\overline{\varphi}, \epsilon, x_0) = +\infty, \; \forall \epsilon \in [0,1).$$

This means that the cost of reducing the starting error may be arbitrarily large for some polynomials from $\mathfrak{J}$ even if x_0 is very close to a solution.

4. TWO THEOREMS

In this section we prove two theorems which play an essential role in the proof of the main result. Although Theorem 4.1 is intuitively obvious its proof is long and difficult. Since this theorem is basic, it would be interesting to find a simpler proof.

We first define a linear operator used below. For any linear information operator $\Re = [L_1, L_2, \ldots, L_n]$, $\Re \in \psi_n$, and any x_0 we define a linear operator $\Re_f$, $f \in H$,

$$(4.1) \quad \Re_f(g) = [L_1(g, z_1), L_2(g, z_2), \ldots, L_n(g, z_n)], \quad \forall g \in H,$$

where $z_1 = x_0$ and

$$(4.2) \quad z_j = z_j(f) = \xi_j(z_1; L_1(f, z_1), L_2(f, z_2), \ldots, L_{j-1}(f, z_{j-1}))$$

are defined by the information operator $\Re$ for the polynomial f. By ker $\Re_f$ we denote the kernel of $\Re_f$. We first establish

Theorem 4.1

For any integer n, any linear information operator $\Re$, $\Re \in \psi_n$, any integer k, any functionals $\varphi_1, \varphi_2, \ldots, \varphi_k \in \Im(\Re)$, and any starting point $x_0 \in \mathcal{L}$, there exists a polynomial $f \in \Im(x_0)$ such that

$$(4.3) \quad x_0, x_1, \ldots, x_k \notin S(f)$$

where $x_i = \varphi_i(x_0; \Re(f, x_0))$, $i = 1, 2, \ldots, k$.   ∎

Proof (induction with respect to n)

We first prove (4.3) for n = 1. Since $\Re \in \psi_1$, there exists a nonzero polynomial h, $h \in H$, satisfying

(4.4) $\mathfrak{N}(h,x_0) = 0$.

Then there exists β, $\beta \in (0,\frac{L}{2})$, such that $h(x_0+\beta) \neq 0$. For positive σ, define

$$f_\sigma(x) = x-x_0-\beta+\sigma h(x).$$

Let $y_1(\sigma),y_2(\sigma),\ldots,y_r(\sigma)$ be the zeros of f_σ where r is the degree of h. From the theory of algebraic functions (see e.g., Wilkinson [63]) we know that $y_1(\sigma) \neq x_0+\beta$ and $y_1(\sigma) \to x_0+\beta$ as σ tends to zero. It is possible to show that the $y_i(\sigma)$ are simple zeros and $|y_i(\sigma)| \to +\infty$ as σ goes to zero, $i \geq 2$. Thus, for sufficiently small σ, $f_\sigma \in \mathfrak{J}(x_0)$ and $f_\sigma(x_0) \neq 0$. Due to (4.4), $\mathfrak{N}(f_\sigma,x_0) = \mathfrak{N}(x-x_0-\beta,x_0)$ which means that

$$x_i = \varphi_i(x_0,\mathfrak{N}(f_\sigma,x_0)) = \varphi_i(x_0,\mathfrak{N}(x-x_0-\beta,x_0))$$

does not depend on σ, $i = 1,2,\ldots,k$. Note that there exists a small σ_1 such that

(4.5) $\{x_0,x_1,\ldots,x_k\} \cap \{y_1(\sigma_1),y_2(\sigma_1),\ldots,y_r(\sigma_1)\} = \emptyset$.

Indeed, for small σ we have $|y_j(\sigma)| > \max\limits_{0\leq i \leq k} |x_i|$ for $j = 2,3,\ldots,r$. Since $y_1(\sigma)$ takes infinitely many values as σ tends to zero, there exists σ_1 such that $y_1(\sigma_1) \neq x_i$, $i = 1,2,\ldots,k$, which proves (4.5). Taking now $f = f_{\sigma_1}$, we get $f \in \mathfrak{J}(x_0)$ and $x_0,x_1,\ldots,x_k \notin S(f)$. This completes the proof of (4.3) for $n = 1$.

Suppose now by induction, that (4.3) holds for $n \leq n_0$. We want to show that (4.3) also holds for $n = n_0+1$. On the contrary assume that there exist $\mathfrak{N}_n^* \in \Psi_n$,

$$\mathfrak{N}_n^* = [L_1^*,L_2^*,\ldots,L_n^*],$$

k, $\omega_1, \omega_2, \ldots, \omega_k \in \mathfrak{z}(\mathfrak{N}_n^*)$ and $x_0 \in \mathbb{C}$ such that for any polynomial $f \in \mathfrak{J}(x_0)$,

(4.6) $\exists i \in [0,k] : x_i(f) = \varphi_i(x_0; \mathfrak{N}_n^*(f,x_0)) \in S(f).$

Define the information operator

$$\mathfrak{N}_{n-1}^*(f,x_0) = [L_1^*(f,z_1), L_2^*(f,z_2), \ldots, L_{n-1}^*(f,z_{n-1})]$$

where z_i are given by $\mathfrak{N}_n^*$. We shall construct functionals $\varphi_1^*, \varphi_2^*, \ldots, \varphi_k^* \in \mathfrak{z}(\mathfrak{N}_{n-1}^*)$ such that the set $\{x_0, x_1^*, x_2^*, \ldots, x_k^*\}$, $x_i^* = x_i^*(f) = \varphi_i^*(x_0; \mathfrak{N}_{n-1}^*(f,x_0))$, contains a zero of f for any polynomial f from $\mathfrak{J}(x_0)$. Since $\mathfrak{N}_{n-1}^* \in \psi_{n_0}$, this will be a contradiction.

Let
$$A_1 = \{f \in \mathfrak{J}(x_0) : f(x_0) = 0\}$$

and let A_2 be the set of all $f \in \mathfrak{J}(x_0)$ for which the functional $L_n^*(\cdot, z_n(f))$ is linearly dependent on the functionals $L_1^*(\cdot, z_1(f))$, $L_2^*(\cdot, z_2(f)), \ldots, L_{n-1}^*(\cdot, z_{n-1}(f))$, i.e., $f \in A_2$ iff there exist constants $c_1, c_2, \ldots, c_{n-1}$ such that

(4.7) $$L_n^*(\cdot, z_n(f)) = \sum_{j=1}^{n-1} c_j L_j^*(\cdot, z_j(f)).$$

Note that c_i depends on the values $z_1(f), z_2(f), \ldots, z_n(f)$ and the functionals $L_1^*, L_2^*, \ldots, L_n^*$. Observe also that for $f \in A_2$ we do not have to compute $L_n^*(f, z_n(f))$ since,

$$L_n^*(f, z_n(f)) = \sum_{j=1}^{n-1} c_j L_j^*(f, z_j(f))$$

is expressed by the previously computed values.

Let

(4.8) $A_3 = \mathfrak{J}(x_0) \setminus (A_1 \cup A_2).$

Then for any $f \in A_3$, $f(x_0) \neq 0$ and

$$(4.9) \quad L_n^*(\cdot, z_n(f)) \notin \text{lin}\{L_1^*(\cdot; z_1(f)), L_2^*(\cdot, z_2(f)), \ldots, L_{n-1}^*(f))\}.$$

For an information operator $\mathfrak{N}$ and $f \in \mathfrak{J}$, let

$$B(\mathfrak{N}_f) = \{\alpha \in \mathbb{C} : \forall h \in \ker \mathfrak{N}_f, \ h(\alpha) = 0\}$$

where $\mathfrak{N}_f$ is a linear operator defined by (4.1). We need the following lemmas.

Lemma 4.1

If $A_3 \neq \emptyset$ then for any $f \in A_3$,

$$S(f) \cap B(\mathfrak{N}_{n-1,f}^*) \neq \emptyset.$$

Proof

From (4.9) there exists a polynomial ζ, $\zeta = \zeta(f) \in H$, such that $L_n^*(\zeta, z_n(f)) = 1$ and $\zeta \in \ker \mathfrak{N}_{n-1,f}^*$. Define

$$g_\sigma(x) = f(x) + \sigma\zeta(x)$$

for $\sigma > 0$. Since f has only simple zeros, then as in the proof for $n = 1$, we can conclude that g_σ has only simple zeros which tend to the zeros of f and to infinity (if the degree of f is less than the degree of ζ) as σ goes to zero. Thus, $g_\sigma \in \mathfrak{J}(x_0)$ for sufficiently small σ. Note that $L_j^*(g_\sigma, z_j(f)) = L_j^*(f, z_j(f))$ for $j = 1, 2, \ldots, n-1$ which means $z_j(f) = z_j(g_\sigma)$ for $j = 1, 2, \ldots, n$. Thus $g_\sigma \notin A_2$. Since $x_0 \notin S(f)$, then x_0 also does not belong to $S(g_\sigma)$ for sufficiently small σ, say $\sigma \in (0, \sigma_0)$. Thus $g_\sigma \in A_3$,

$$\exists i_0 = i_0(\sigma) \in [1, k] : x_{i_0}(g_\sigma) = \varphi_{i_0}(x_0; \mathfrak{N}_n^*(g_\sigma, x_0)) \in S(g_\sigma), \quad \forall \sigma \in (0, \sigma_0).$$

Let h be an arbitrary polynomial, $h \in \ker \pi_{n,f}^*$. Consider

$$g_{\sigma,\beta}(x) = g_\sigma(x) + \beta h(x)$$

where β is a sufficiently small number. Then $g_{\sigma,\beta} \in A_3$ and $\pi_n^*(g_\sigma, x_0) = \pi_n^*(g_{\sigma,\beta}, x_0)$ which means that $x_i(g_{\sigma,\beta}) = x_i(g_\sigma)$ does not depend on β, $i = 1, 2, \ldots, k$, and therefore

$$h(x_{i_0}(g_\sigma)) = 0, \quad \forall \sigma \in (0, \sigma_0).$$

Since h is arbitrary, this yields that $S(g + \sigma\zeta) \cap B(\pi_{n,f}^*)$ is nonempty.

Let $h \in \ker \pi_{n,f}^*$ be a nonzero polynomial. Then there exists $i_0 = i_0(\sigma) \in [1,k]$ such that x_{i_0} is a zero of $g + \sigma\zeta$ and h, $\sigma \in (0, \sigma_0)$. Since σ takes infinitely many values, there exist distinct σ_1 and σ_2, both from $(0, \sigma_0)$, such that $i_0(\sigma_1) = i_0(\sigma_2)$. Let $i = i_0(\sigma_1) = i_0(\sigma_2)$ and $x_i^* = x_i(g_{\sigma_1}) = x_i(g_{\sigma_2})$. Then

$$0 = g_{\sigma_1}(x_i^*) = f(x_i^*) + \sigma_1\zeta(x_i^*)$$
$$0 = g_{\sigma_2}(x_i^*) = f(x_i^*) + \sigma_2\zeta(x_i^*).$$

This means $f(x_i^*) = \zeta(x_i^*) = 0$ and since h is arbitrary, we get

$$x_i^* \in S(f) \cap S(\zeta) \cap B(\pi_{n,f}^*).$$

Since $\ker \pi_{n-1,f}^* = \lim\{\zeta\} \oplus \ker \pi_{n,f}^*$, see (4.9) and the definition of ζ, we get $x_i^* \in S(f) \cap B(\pi_{n-1,f}^*)$. Thus $S(f) \cap B(\pi_{n-1,f}^*)$ is nonempty which completes the proof of Lemma 4.1. ■

For any f, $f \in \mathfrak{J}(x_0)$, let $r = r(f)$, $j = j(f)$, $(j = 1, 2, \ldots r)$ be indices such that $\{L_{j_1}^*(\cdot, z_{j_1}(f)), L_{j_2}^*(\cdot, z_{j_2}(f)), \ldots, L_{j_r}^*(\cdot, z_{j_r}(f))\}$ is a basis of the space $\lim\{L_1^*(\cdot, z_1(f)), L_2^*(\cdot, z_2(f)), \ldots, L_{n-1}^*(f))\}$. Let $\zeta_1, \zeta_2, \ldots, \zeta_r$, $\zeta_i(f) \in H$, be polynomials satisfying

$$L^*_{j_s}(\zeta_i, z_{j_s}(f)) = \begin{cases} 1 & \text{if } s = i, \\ \\ 0 & \text{if } s \neq i. \end{cases}$$

We define

$$w_f = \sum_{s=1}^{r} L^*_{j_s}(f, z_{j_s}(f)) \zeta_s$$

and

$$A_4 = \{f \in \mathfrak{J}(x_0) : S(w_f) \cap B(\mathfrak{N}^*_{n-1,f}) \neq \emptyset\}.$$

Lemma 4.2

(i) $A_3 \subset A_4$,

(ii) if $A_4 \neq \emptyset$ then for any $f \in A_4$,

$$S(w_f) \cap B(\mathfrak{N}^*_{n-1,f}) \subset S(f).$$

Proof

Without loss of generality we can assume that A_3 is nonempty. Let $f \in A_3$ be arbitrary. Then $h_f \overset{df}{=} f - w_f \in \ker \mathfrak{N}^*_{n-1,f}$ and from Lemma 4.1, there exists $\alpha \in S(f) \cap B(\mathfrak{N}^*_{n-1,f})$. Thus, $w_f(\alpha) = f(\alpha) - h_f(\alpha) = 0$ which means that $S(w_f) \cap B(\mathfrak{N}^*_{n-1,f})$ is nonempty. Thus, $f \in A_4$ which proves that $A_3 \subset A_4$.

To prove (ii), let f be an arbitrary polynomial from A_4. There exists $\alpha_1 \in S(w_f) \cap B(\mathfrak{N}^*_{n-1,f})$. Since $h_f \in \ker \mathfrak{N}^*_{n-1,f}$, $f(\alpha_1) = w_f(\alpha_1) + h_f(\alpha_1) = 0$ which means that $\alpha_1 \in S(f)$. Thus, Lemma 4.2 is proven. ∎

Note that knowing $\mathfrak{N}^*_{n-1}(f, x_0)$ we can verify whether f belongs to A_i, $i = 2,4$. Furthermore for any $\tilde{f} \in \mathfrak{J}(x_0)$ with $\mathfrak{N}^*_{n-1}(\tilde{f}, x_0) = \mathfrak{N}^*_{n-1}(f, x_0)$, $\tilde{f} \in A_i$ iff $f \in A_i$, $i = 2,4$. For $i = 1,2,\ldots,k$, define

$$
\varphi_i^*(x_0; \mathfrak{N}_{n-1}^*(f,x_0)) = \begin{cases} \alpha \in S(w_f) \cap B(\mathfrak{N}_{n-1,f}^*) & \text{if } \tilde{f} \in A_4, \\[4mm] \varphi_i(x_0; [\mathfrak{N}_{n-1}^*(f,x_0), \overset{n-1}{\underset{j=1}{\sum}} c_j L_j^*(f,z_j(f))]) & \text{if } \tilde{f} \in A_2 \backslash A_4, \\[4mm] 0 & \text{otherwise,} \end{cases}
$$

where $\tilde{f} \in \mathfrak{J}(x_0)$ and $\mathfrak{N}_{n-1}^*(\tilde{f},x_0) = \mathfrak{N}_{n-1}^*(f,x_0)$. Thus, the functionals φ_i^* are well-defined and $\varphi_i^* \in \Phi(\mathfrak{N}_{n-1}^*)$. Furthermore

(4.10) $\forall f \in \mathfrak{J}(x_0), \exists i_0 \in [0,k] : x_{i_0} \in S(f)$.

Indeed, if $f \in A_4$ then, due to Lemma 4.2(ii), $\varphi_i^*(x_0, \mathfrak{N}_{n-1}^*(f,x_0)) = \alpha \in S(f)$ for any $i = 1,2,\ldots,k$. If $f \in \mathfrak{J}(x_0)\backslash A_4$, then due to Lemma 4.2(i), $f \in A_1 \cup A_2$ which means that either $x_0 \in S(f)$ or $\varphi_{i_0}^*(x_0; \mathfrak{N}_{n-1}^*(f,x_0)) = \varphi_{i_0}(x_0; \mathfrak{N}_n^*(f,x_0)) \in S(f)$ for some $i_0 \in [1,k]$. Since $\varphi_i^* \in \Phi(\mathfrak{N}_{n-1}^*)$ and $\mathfrak{N}_{n-1}^* \in \psi_{n_0}$, (4.10) contradicts the inductive assumption. Thus, the proof of Theorem 4.1 is completed. ▨

Theorem 4.1 says that for any linear information operator $\mathfrak{N}$ and any finite number of functionals $\varphi_1, \varphi_2, \ldots, \varphi_k \in \Phi(\mathfrak{N})$, there exists a polynomial $f \in \mathfrak{J}(x_0)$ such that no point $x_i = \varphi_i(x_0; \mathfrak{N}(f,x_0))$ is a zero α of f. Now we show that x_i cannot approximate α better than x_0.

Theorem 4.2

For any integer n, any linear information operator $\mathfrak{N}$, $\mathfrak{N} = [L_1, L_2, \ldots, L_n] \in \psi_n$, any integer k, any functionals $\varphi_1, \varphi_2, \ldots, \varphi_k \in \Phi(\mathfrak{N})$ and any starting point $x_0 \in \mathbb{C}$, there exists a polynomial $f \in \mathfrak{J}(x_0)$ such that

$$
\underset{i=0,1,\ldots,k}{\min} \text{dist}(x_i, S(f)) = \text{dist}(x_0, S(f)) \neq 0
$$

where $x_i = x_i(f) = \varphi_i(x_0; \mathfrak{N}(f,x_0))$. ■

Proof

From Theorem 4.1, there exists a polynomial g, $g \in \mathfrak{I}(x_0)$, such that $x_0, x_1 = x_1(g), \ldots, x_k = x_k(g) \notin S(g)$. Let $I = \{i \in [1,k] : x_i(g) \neq x_0\}$. If $I = \emptyset$ then for $f = g$ we have

$$0 \neq \text{dist}(x_0, S(f)) = \text{dist}(x_i(f), S(f)), \quad \forall i = 1, 2, \ldots, k,$$

which completes the proof.

Suppose therefore that $I \neq \emptyset$. Consider a polynomial w of the form

$$(4.11) \quad w(x) = \prod_{i \in I}(x - x_i)^m (x - x_0) \sum_{j=0}^{n} a_j x^j, \quad m = \max\{3n, \deg g\},$$

satisfying

$$(4.12) \quad \mathfrak{R}_g(w) = 0.$$

Note that (4.12) is equivalent to the following system of n homogeneous linear equations

$$(4.13) \quad \sum_{j=0}^{n} a_j L_s \left(\prod_{i \in I}(x - x_i)^m (x - x_0) x^j, z_s(g) \right) = 0 \quad \text{for} \quad s = 1, 2, \ldots, n.$$

Since (4.13) has more unknowns than equations, there exists a non-zero polynomial satisfying (4.11) and (4.12). Consider the factorization of w,

$$w(x) = (x - x_0)^{p_0} \prod_{i \in I}(x - x_i)^{p_i} \prod_{j=1}^{r}(x - y_j)^{s_j}$$

for some r, $r \leq n$, $s_1, s_2, \ldots, s_r$ and p_0, p_i for $i \in I$ where $y_j \neq x_i$ for any i and j. Due to (4.11),

$$(4.14) \quad p_0 \leq n+1 \text{ and } p_i \geq 3n \text{ for } i \in I.$$

For $\sigma > 0$ define

$$f_\sigma(x) = g(x) + \frac{1}{\sigma}w(x).$$

Since g has only simple zeros then for sufficiently small σ, $f_\sigma \in \mathfrak{J}$. Furthermore from the theory of algebraic functions, see e.g., Wilkinson [63], f_σ has zeros $x_{i_j}(\sigma)$, $i \in I \cup \{0\}$, $j = 1,2,\ldots,p_i$, satisfying

$$(4.15) \quad |x_{i_j}(\sigma) - x_i| = \left| \frac{p_i! g(x_i)}{w^{(p_i)}(x_i)} \right|^{1/p_i} \sigma^{1/p_i}(1+o(1)) \overset{df}{=} M_i \sigma^{1/p_i}(1+o(1))$$

and zeros $y_{i_j}(\sigma)$, $i = 1,2,\ldots,r$ and $j = 1,2,\ldots,s_i$, which tend to y_i when σ goes to zero. Thus, $f_\sigma \in \mathfrak{J}(x_0)$ for sufficiently small σ. Since $\mathfrak{N}(f_\sigma,x_0) = \mathfrak{N}(g,x_0)$ then $x_i(f_\sigma) = x_i(g)$ for $i = 1,2,\ldots,k$, and therefore

$$\text{dist}(x_i(f_\sigma),S(f_\sigma)) = M_i \sigma^{1/p_i}(1+o(1)), \quad \text{for } i = 0 \text{ and } i \in I.$$

From this and (4.15) we get for $i \in I$,

$$\frac{\text{dist}(x_i(f_\sigma),S(f_\sigma))}{\text{dist}(x_0,S(f))} = \frac{M_i}{M_0} \sigma^{1/p_i - 1/p_0} = \frac{M_i}{M_0} \sigma^{\frac{1-2n}{3n(n+1)}}(1+o(1)).$$

Since M_i are bounded away from zero, there exists $\sigma_0, \sigma_0 > 0$, such that

$$\frac{M_i}{M_0} \sigma_0^{\frac{1-2n}{3n(n+1)}}(1+o(1)) \geq 1, \quad \forall i \in I.$$

This implies $\text{dist}(x_i(f_{\sigma_0}),S(f_{\sigma_0})) \geq \text{dist}(x_0,S(f_{\sigma_0}))$, $\forall i \in I$. Note that $\text{dist}(x_i(f_{\sigma_0}),S(f_{\sigma_0})) = \text{dist}(x_0,S(f_{\sigma_0})) \neq 0$, $\forall i \notin I$. Thus putting $f = f_{\sigma_0}$, we get $f \in \mathfrak{J}(x_0)$ and

$$\min_{i=0,1,\ldots,k} \text{dist}(x_i(f),S(f)) = \text{dist}(x_0,S(f)) \neq 0$$

which completes the proof. ∎

5. MAIN RESULT

We are ready to prove the main result of this paper.

Theorem 5.1

For any positive L, any sequence of linear information operators $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$, any iteration without memory $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}(\overline{\mathfrak{N}})$, and any starting point $x_0 \in \mathbb{C}$,

$$N(\overline{\varphi}, \epsilon, x_0) = +\infty, \quad \forall \epsilon < 1.$$

Proof

Suppose on the contrary that for some $\epsilon < 1$, $\overline{\mathfrak{N}}$, $\overline{\varphi}$ and x_0, $k = N(\overline{\varphi}, \epsilon, x_0)$ is finite. This means that

(5.1) $\forall f \in \mathfrak{I}(x_0)$, $\exists i_0 \in [1,k] : \text{dist}(x_{i_0}, S(f)) \leq \epsilon \, \text{dist}(x_0, S(f))$

where $x_i = \varphi_i(x_0; \mathfrak{N}_i(f, x_0))$. Consider now the operator defined by

(5.2) $\mathfrak{N}^*(f, x_0) = [\mathfrak{N}_1(f, x_0), \mathfrak{N}_2(f, x_0), \ldots, \mathfrak{N}_k(f, x_0)]$.

Of course, $\mathfrak{N}^*$ is a linear information operator in the sense of Definition (2.2). Define functionals $\varphi_i^* \in \Phi(\mathfrak{N}^*)$,

(5.3) $\varphi_i^*(x_0; \mathfrak{N}^*(f, x_0)) = \varphi_i(x_0; \mathfrak{N}_i(f, x_0))$.

From Theorem 4.2 we know that for functionals φ_i^* there exists a polynomial f_0, $f_0 \in \mathfrak{I}(x_0)$, such that

(5.4) $\text{dist}(x_i^*, S(f_0)) \geq \text{dist}(x_0, S(f_0)) \neq 0$, $\forall i = 1, 2, \ldots, k$,

where $x_i^* = x_i^*(f_0) = \varphi_i^*(x_0; \mathfrak{N}^*(f_0, x_i))$, $i = 1, 2, \ldots, k$. Due to (5.3), $x_i^*(f_0) = x_i(f_0)$. From (5.1) there exists $i_0 = i_0(f_0)$ such that

$$\text{dist}(x_{1_0}, S(f_0)) \leq \epsilon \ \text{dist}(x_0, S(f_0)) < \text{dist}(x_0, S(f_0))$$

which contradicts (5.4). Hence Theorem 5.1 is proven. ■

From Theorem 5.1 and (3.8) follows

Corollary 5.1

For any positive L, any sequence of linear information operators $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$, any iteration without memory $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}(\overline{\mathfrak{N}})$, and any starting point $x_0 \in \mathbb{C}$,

$$\text{comp}(\overline{\varphi}, \epsilon, x_0) = +\infty, \quad \forall \epsilon < 1. \qquad ■$$

6. MEMORY

In this section we briefly extend all previous definitions and results to iterations with memory. Let H and $\mathfrak{J}$ be the class of polynomials defined as in Section 2. Let m, m > 0, be an integer, and let $L_i : H \times \mathbb{C}^{m+i} \to \mathbb{C}$ be a functional which is linear with respect to f. Then the <u>linear information operator</u> $\mathfrak{N}$, $\mathfrak{N} = [L_1, L_2, \ldots, L_n] : H \times \mathbb{C}^{m+1} \to \mathbb{C}^n$, is defined as

$$(6.1) \quad \mathfrak{N}(f, x_0, x_{-1}, \ldots, x_{-m}) = [L_1(f, z_1, z_2, \ldots, z_{m+1}), L_2(f, z_1, z_2, \ldots, z_{m+2}), \ldots,$$
$$L_n(f, z_1, z_2, \ldots, z_{m+n})]$$

where $z_1 = x_0, z_2 = x_{-1}, \ldots, z_{m+1} = x_{-m}$ and $z_j = \xi_j(z_1, \ldots, z_{m+1}, L_1(f, z_1, z_2, \ldots, z_{m+1}),$
$L_2(f, z_1, z_2, \ldots, z_{m+2}), \ldots, L_{j-1}(f, z_1, z_2, \ldots, z_{m+j-1}))$ for some functions ξ_j,
$j = m+2, m+3, \ldots, n$. Let $\psi_{n,m}$ be the class of all such information operators.
Let $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$ be a sequence of $\mathfrak{N}_i$, $\mathfrak{N}_i \in \psi_{n_i, m}$. For distinct $x_0, x_{-1}, \ldots, x_{-m}$ we construct a sequence of approximations $\{x_i\}$ by the formula

$$(6.2) \quad x_i = \varphi_i(x_0, x_{-1}, \ldots, x_{-m}; \mathfrak{N}_i(f, x_0, x_{-1}, \ldots, x_{-m}))$$

where $\varphi_i = D_{\varphi_i} \subset \mathbb{C}^{1+m+n_i} \to \mathbb{C}$ are functionals, $\varphi_i \in \Phi_m(\mathfrak{N}_i)$. Let now $\overline{\varphi} = \{\varphi_i\}$ be a sequence of functionals φ_i, $\varphi_i \in \Phi_m(\mathfrak{N}_i)$. Then $\overline{\varphi}$ is called an <u>iteration</u> <u>with memory</u> using $\overline{\mathfrak{N}}$, $\overline{\varphi} \in \overline{\Phi}_m(\overline{\mathfrak{N}})$.

Let $\overline{\mathfrak{N}} = \{\mathfrak{N}_i\}$ be a sequence of linear information operators with memory, $\mathfrak{N}_i \in \psi_{n_i, m}$, $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}_m$, ϵ be a positive number, and $x_0, x_{-1}, \ldots, x_{-m}$ be given distinct points. For any $f \in \mathfrak{J}$, define

$$(6.3) \quad N = N(\overline{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$$

as the minimal integer, if it exists, such that

(6.4) $\text{dist}(x_N, S(f)) \le \epsilon \, \text{dist}(x_0, S(f))$,

and $N = +\infty$ otherwise. Let $\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$ be the cost of computing x_N. Let L, L > 0, be a given constant. Then

(6.5) $N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \overset{\text{df}}{=} \underset{f \in \mathfrak{J}(x_0)}{\sup} N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$

where $\mathfrak{J}(x_0)$ is defined by (3.6). Similarly, let

(6.6) $\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \overset{\text{df}}{=} \underset{f \in \mathfrak{J}(x_0)}{\sup} \text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}, f)$.

As before, there exists a positive c such that

(6.7) $\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) \ge cN(\bar{\varphi}, \epsilon, x_0, x_1, \ldots, x_{-m})$, $\forall \bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}$.

By a technique similar to the proof of Theorem 5.1 it is possible to prove

Theorem 6.1

For any positive L, any m, m > 0, any sequence of linear information operators with memory $\bar{\mathfrak{N}}$, any iteration $\bar{\varphi}$, $\bar{\varphi} \in \Phi_m(\bar{\mathfrak{N}})$ and any distinct starting points $x_0, x_{-1}, \ldots, x_{-m} \in \mathbb{C}$

$\text{comp}(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) = N(\bar{\varphi}, \epsilon, x_0, x_{-1}, \ldots, x_{-m}) = +\infty$, $\forall \epsilon < 1$. ∎

Remark 6.1

In practice one often wants to reduce a residual error, i.e., to find a point x_k such that

(6.8) $|f(x_k)| \le \epsilon |f(x_0)|$

for some $\epsilon < 1$. Note that (6.8) does not imply that x_k is closer to a solution $\alpha \in S(f)$ than x_0. For problem (6.8) we pose the same question:

What is the complexity of any iteration $\bar{\varphi}$ using linear information solving (6.8)?

It can be proven that the complexity of $\bar{\varphi}$ is infinity even for a subclass of $\mathfrak{J}(x_0)$. More precisely, let T be a finite dimensional linear operator which maps H onto $\mathbb{C}^s$. Let $\vec{c} \in \mathbb{C}^s$. Define

$$\mathfrak{J}(x_0, T, \vec{c}) = \{f \in \mathfrak{J}(x_0) : T(f) = \vec{c}\}.$$

For any iteration with memory $\bar{\varphi} \in \bar{\Phi}_m(\bar{\mathfrak{N}})$, any $\epsilon < 1$, any $x_0 \in \mathbb{C}$, and any $f \in \mathfrak{J}(x_0, T, \vec{c})$, define $N' = N'(\bar{\varphi}, \epsilon, \vec{c}, x_0, f)$ as the minimal integer, if it exists, such that $|f(x_{N'})| \leq \epsilon |f(x_0)|$, and $N' = +\infty$ otherwise. Furthermore, let

$$N'(\bar{\varphi}, \epsilon, \vec{c}, x_0) = \sup_{f \in \mathfrak{J}(x_0, T, \vec{c})} N'(\bar{\varphi}, \epsilon, \vec{c}, x_0, f).$$

Then for any sequence of linear information operators with memory $\bar{\mathfrak{N}}$, any $\bar{\varphi} \in \Phi_m(\bar{\mathfrak{N}})$, any T, any nonzero $\vec{c} \in \mathbb{C}^s$, and any distinct starting points $x_0, x_{-1}, \ldots, x_{-m} \in \mathbb{C}$, we have

$$N'(\bar{\varphi}, \epsilon, \vec{c}, x_0) = +\infty, \quad \forall \epsilon < 1.$$

In particular, this also holds for $T(f) = f(x_0)$, i.e., for the class of all complex polynomials from $\mathfrak{J}(x_0)$ which have a fixed value at x_0, $f(x_0) = c \neq 0$.

We want to stress that Theorems 5.1 and 6.1 can also be proven for the class $\mathfrak{J}(x_0, T, \vec{c})$. The proof, however, is more complicated. ∎

7. OPEN PROBLEMS

In this section we pose a number of open problems which are relevant to the questions studied in this paper.

In Theorem 6.1 we prove that for any $m \geq 0$, any linear information $\overline{\pi} = \{\overline{\pi}_i\}$, $\pi_i \in \Psi_{n_i,m}$, any iteration $\overline{\varphi} = \{\varphi_i\} \in \overline{\Phi}_m(\overline{\pi})$ and any integer k, there exists a "difficult" polynomial f, $f \in \mathfrak{I}(x_0)$, i.e., a polynomial which requires at least k+1 iterative steps to reduce the starting error $\text{dist}(x_0, S(f))$. Let $P = P(\overline{\pi}, \overline{\varphi}, k)$ be the set of all such difficult polynomials and let $d = d(\overline{\pi}, \overline{\varpi}, k)$ be the minimal degree of such polynomials, i.e.,

$$d \stackrel{\mathrm{df}}{=} \min_{f \in P} \deg f.$$

Problem 1

Find d as a function of m, k, and $n_1, n_2, \ldots, n_k$. ∎

It can be shown that

$$(7.1) \quad d \leq (k+2)\left(2 + \sum_{i=1}^{k} n_i\right) + k.$$

In general, this bound is not sharp. For instance, for a <u>stationary iteration</u>,

$$(7.2) \quad d \leq (k+1)(n_1+1) + k.$$

By a stationary iteration we mean an iteration which constructs a sequence of approximations by the formula

$$(7.3) \quad x_{i+1} = \varphi_1(x_i, x_{i-1}, \ldots, x_{i-m}; \pi_1(f, x_i, x_{i-1}, \ldots, x_{i-m}))$$

for some $\pi_1 \in \Psi_{n_1,m}$ and $\varphi_1 \in \Phi(\pi_1)$.

Theorem 6.1 states that the set P of "difficult" polynomials is nonempty. It would also be interesting to investigate how "large" this set is. We therefore pose

Problem 2

What proportion of polynomials from $\mathfrak{J}(x_0)$ is "difficult"? ∎

In this paper we have restricted ourselves to the class $\mathfrak{J}$ of all complex polynomials having only simple zeros. It is interesting to find for which classes of problems the same negative result holds, i.e., for which classes of problems the complexity of finding x_N using linear information is infinite. Does this hold for the class of all real polynomials with all simple zeros? We summarize this in

Problem 3

Characterize the classes of problems for which the complexity of finding x_N is infinite. ∎

From our results it follows that to make the complexity of finding x_N finite, it is necessary to use some nonlinear information about f. An important open problem is to characterize nonlinear information which yields not only finite but relatively small complexity. On the other hand, there exists nonlinear information for which the complexity is still infinity. For instance, if $\mathfrak{N}(f,x_0) = G(L_1(f,z_1),L_2(f,z_2),\ldots,L_n(f,z_n))$ where $L_1,\ldots,L_n$ are, as always, linear functionals with respect to f and G is an arbitrary operator (nonlinear in general), then it is obvious that for any sequence $\overline{\mathfrak{N}}$ of such iteration operators, the complexity of finding x_N is infinite. We propose

Problem 4

(i) For a given nonincreasing function g, $g : [0,1) \to R$, find information $\bar{\mathfrak{N}}$ and an iteration $\bar{\varphi}$ using $\bar{\mathfrak{N}}$ such that the complexity $comp(\bar{\varphi}, \epsilon, x_0) \leq g(\epsilon)$ for any $\epsilon \in [0,1)$.

(ii) Characterize the class of all information for which the complexity of finding x_N is finite. ∎

ACKNOWLEDGMENT

REFERENCES

Wasilkowski [78]   Wasilkowski, G. W., "Can Any Stationary Itera-
                          tion Using Linear Information be Globally Con-
                          vergent?," Department of Computer Science
                          Report, Carnegie-Mellon University, 1978. To
                          appear in JACM.

Wasilkowski [79]   Wasilkowski, G. W., "The Strength of Nonsta-
                          tionary Iteration," Department of Computer
                          Science Report, Carnegie-Mellon University, 1979.

Wilkinson [63]   Wilkinson, J. H., Rounding Errors in Algebraic
                          Processes, Prentice-Hall, Inc., 1963.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CMU-CS-79-139	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ANY ITERATION FOR POLYNOMIAL EQUATIONS USING LINEAR INFORMATION HAS INFINITE COMPLEXITY		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G. W. WASILKOWSKI		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0370
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Department Pittsburgh, PA 15213		10. PROGRAM ELEMENT. PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE September 1979
		13. NUMBER OF PAGES 30 Pages
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)